

# Constructing Boolean Functions With Potential Optimal Algebraic Immunity Based on Additive Decompositions of Finite Fields

Baofeng Wu

State Key Lab of Information Security  
Institute of Information Engineering  
Chinese Academy of Sciences  
Beijing 100093, China  
Email: wubaofeng@iie.ac.cn

Qingfang Jin and Zhuojun Liu

Key Lab of Mathematics Mechanization  
Academy of Mathematics and Systems Science  
Chinese Academy of Sciences  
Beijing 100190, China  
Email: qfjin@amss.ac.cn  
and zliu@mmrc.iss.ac.cn

Dongdai Lin

State Key Lab of Information Security  
Institute of Information Engineering  
Chinese Academy of Sciences  
Beijing 100093, China  
Email: ddlin@iie.ac.cn

**Abstract**—We propose a general approach to construct cryptographic significant Boolean functions of  $(r+1)m$  variables based on the additive decomposition  $\mathbb{F}_{2^{rm}} \times \mathbb{F}_{2^m}$  of the finite field  $\mathbb{F}_{2^{(r+1)m}}$ , where  $r$  is odd and  $m \geq 3$ . A class of unbalanced functions are constructed first via this approach, which coincides with a variant of the unbalanced class of generalized Tu-Deng functions in the case  $r = 1$ . This class of functions have high algebraic degree, but their algebraic immunity does not exceeds  $m$ , which is impossible to be optimal when  $r > 1$ . By modifying these unbalanced functions, we obtain a class of balanced functions which have optimal algebraic degree and high nonlinearity (shown by a lower bound we prove). These functions have optimal algebraic immunity provided a combinatorial conjecture on binary strings which generalizes the Tu-Deng conjecture is true. Computer investigations show that, at least for small values of number of variables, functions from this class also behave well against fast algebraic attacks.

## I. INTRODUCTION

Constructing Boolean functions satisfying all main criteria has attracted a lot of attention in recent year. Among all these criteria, optimal algebraic immunity seems necessary due to the great success of algebraic attacks introduced (improved, more definitely) by Courtious and Meier to some well-known Boolean-function-based stream ciphers [5]. Other criteria for Boolean functions that can play as potential candidates in designing such LFSR-based pseudo-random generators as filter generators include balancedness, high algebraic degree and high nonlinearity. Besides, because of the existence of the improved algebraic attacks, the fast algebraic attacks (FAA's) [6], a good behavior against FAA's is also required for Boolean functions to be usable in cryptography.

In fact, it is a big challenge to construct Boolean function with optimal algebraic immunity together with all other good cryptographic properties and there has been little work on such a topic until 2008. In their pioneering work [3], Carlet and Feng constructed a classes of balanced functions with optimal algebraic immunity, optimal algebraic degree, high nonlinearity and good behavior against FAA's (verified by computers initially in [3] and confirmed by Liu et al. in

[10] theoretically very recently). Their construction is based on finite fields and the proof of optimal immunity of the constructed functions is mostly based on univariate representations of Boolean functions. Motivated by their idea of construction, Tu and Deng went a further step. They constructed a class of balanced functions of even number of variables with optimal algebraic degree, high nonlinearity and potential optimal algebraic immunity. By “potential” we mean that the optimal algebraic immunity is up to a conjecture on binary strings (known as the Tu-Deng conjecture now) which is not mathematically proved. In fact, their functions are modified from functions belonging to a subclass of the well-known  $\mathcal{PS}_{\text{ap}}$  class of bent functions. A weakness of this class of functions is their immunity against FAA's is bad [1]. However, the idea of Tu and Deng's construction is enlightening. Adopting similarly techniques, Tang et al. constructed a class of functions satisfying all main criteria. It is remarkable that the optimal algebraic immunity of this class of functions is based on a combinatorial fact firstly conjectured by Tang et al. and proved by Cohen and Flori [4] afterwards. Based on a general conjecture involving a parameter which can be chosen rather freely mentioned in [14] (known as the generalized Tu-Deng conjecture), Jin et al. proposed a construction of Boolean functions with optimal immunity covering those in [15] and [14]. All the functions obtained in [15], [14], [8] are constructed from a decomposition of the finite field into a direct sum of a subfield and a copy of it, and the proofs of (potential) optimal algebraic immunity of them are mostly based on the so-called bivariate representations of Boolean functions.

Note that the decompositions of finite fields used in [15], [14], [8] are all additive ones. More precisely, the additive group of a finite field is decomposed into a direct sum of two additive groups with equal sizes to construct functions. Therefore, to generalize the constructions in [15], [14], [8], a natural idea is to use decompositions of additive groups of finite fields into direct sums of additive groups with unequal

sizes. Besides, to study properties of functions constructed from such kinds of decompositions, the summands of a decomposition are preferred both to be additive groups of certain finite fields.

In the present paper, we devote to realize this idea. By decomposing the additive group of the finite field  $\mathbb{F}_{2^{(r+1)m}}$  into a direct sum of additive groups of the finite fields  $\mathbb{F}_{2^{rm}}$  and  $\mathbb{F}_{2^m}$  for an odd integer  $r \geq 1$  and an integer  $m \geq 3$ , we construct a class of  $(r+1)m$ -variable unbalanced Boolean functions in a similar manner with those in [15], [14], [8]. This class coincides with a variant of the unbalanced class proposed in [8] when  $r = 1$ , but when  $r > 1$ , some properties of functions belonging to it are different, say, their algebraic immunity will never be optimal. However, after a modification of this class, we obtain a class of balanced functions with optimal algebraic immunity provided a combinatorial conjecture is true, but the proof of optimal algebraic immunity of these functions in the case  $r > 1$  is quite different from the proof in the case  $r = 1$ , i.e. the proof of optimal algebraic immunity of the balanced functions obtained in [8]. In fact, in the case  $r > 1$ , the first things that should be made clear are, how to represent functions defined from the additive decomposition before-mentioned and how to study properties of such functions under this kind of representation if we can find it.

The rest of the paper is organized as follows. In the following section, we recall some basic notions about Boolean functions and talk about bivariate representations of Boolean functions over direct sums of finite fields. In Section III, we present a general combinatorial conjecture on binary strings. In Section IV, we propose a class of unbalanced functions to make our idea of constructing a class of balanced functions with good cryptographic properties, which is proposed in Section V, more clear. Concluding remarks are given in Section VI.

## II. PRELIMINARIES

In this section, we provide some basic notations and facts about Boolean functions. For more details, we refer to [2].

### A. Boolean functions and related basic notions

Let  $\mathbb{F}_2$  be the binary finite field and  $\mathbb{F}_2^n$  be the  $n$ -dimensional vector space over  $\mathbb{F}_2$ . Any mapping from  $\mathbb{F}_2^n$  to  $\mathbb{F}_2$  is called an  $n$ -variable Boolean function. Obviously, the set  $\mathbb{B}_n$  consisting of all  $n$ -variable Boolean functions forms an  $\mathbb{F}_2$ -algebra of dimension  $2^n$ . For a Boolean function  $f \in \mathbb{B}_n$ , its support is defined as

$$\text{supp}(f) = \{x \in \mathbb{F}_2^n \mid f(x) = 1\},$$

and the cardinality of this set, denoted by  $\text{wt}(f)$ , is called its Hamming weight.  $f$  is called balanced if  $\text{wt}(f) = 2^{n-1}$ . Furthermore, for another Boolean function  $g \in \mathbb{B}_n$ , the distance between  $f$  and  $g$  is defined as  $d(f, g) = \text{wt}(f + g)$ . Abusing notations, we also denote the Hamming weight of a vector  $v \in \mathbb{F}_2^n$ , i.e. the number of nonzero positions of  $v$ , to be  $\text{wt}(v)$ . Besides, for an integer  $i$ , we denote by  $\text{wt}_n(i)$  the number of 1's in the binary expansion of the

reduction of  $i$  modulo  $(2^n - 1)$  in the complete residue system  $\{0, 1, \dots, 2^n - 2\}$ . Obviously,  $\text{wt}_n(-u) = n - \text{wt}_n(u)$  when  $2^n - 1 \nmid u$ .

By Lagrange interpolation, every  $n$ -variable Boolean function  $f$  can be uniquely represented as

$$f(x_1, \dots, x_n) = \sum_{I \subseteq \{1, 2, \dots, n\}} a_I \prod_{i \in I} x_i, \quad a_I \in \mathbb{F}_2.$$

The deep reason for the existence of such kinds of representations of Boolean functions, often known as algebraic normal forms (ANF's) of Boolean functions, lies in the isomorphism between  $\mathbb{F}_2$ -algebras

$$\mathbb{B}_n \cong \mathbb{F}_2[x_1, x_2, \dots, x_n] / \langle x_1^2 + x_1, \dots, x_n^2 + x_n \rangle.$$

Thanks to its ANF, we can define the algebraic degree of  $f$ ,  $\deg f$ , to be the degree of  $f(x_1, \dots, x_n)$  as a multivariate polynomial, i.e.  $\deg f = \max_{I \subseteq \{1, 2, \dots, n\}} \{|I| \mid a_I \neq 0\}$ . Boolean functions of degree at most 1 are called affine functions. The minimum distance between  $f$  and all affine functions is called the nonlinearity of  $f$  and denoted to be  $\mathcal{N}_f$ . This notion characterizes how different is  $f$  from the simplest Boolean functions.

As is well known that the additive group of the finite field  $\mathbb{F}_{2^n}$  is an  $n$ -dimensional vector space over  $\mathbb{F}_2$ , hence by Lagrange interpolation, the Boolean function  $f$  can also be represented by a univariate polynomial over  $\mathbb{F}_{2^n}$  of the form

$$f(x) = \sum_{i=0}^{2^n-1} f_i x^i.$$

However, since  $f$  satisfies  $f^2(x) = f(x)$  for any  $x \in \mathbb{F}_{2^n}$ , there are some restrictions on the coefficients of this kind of univariate representation. This kind of representation implies that as  $\mathbb{F}_2$ -algebras,  $\mathbb{B}_n$  can be viewed as a subalgebra of  $\mathbb{F}_{2^n} / \langle x^{2^n} + x \rangle$ . Comparing dimensions, we can also obtain the isomorphism

$$\mathbb{F}_{2^n} / \langle x^{2^n} + x \rangle \cong \mathbb{B}_n \otimes_{\mathbb{F}_2} \mathbb{F}_{2^n}.$$

It can be deduced that, under its univariate representation, the algebraic degree of  $f$  is in fact

$$\deg f = \max_{0 \leq i \leq 2^n-1} \{\text{wt}_n(i) \mid f_i \neq 0\}.$$

### B. Bivariate representations of Boolean functions

In fact, representations of Boolean functions are more flexible than what can be fully described. In this subsection, we introduce the bivariate representations of Boolean functions, which have already been mentioned in [11] without explaining details.

Assume  $n = n_1 + n_2$  for two integers  $n_1, n_2 \geq 1$ . We can decompose the additive group of  $\mathbb{F}_{2^n}$  into a direct sum of additive groups of  $\mathbb{F}_{2^{n_1}}$  and  $\mathbb{F}_{2^{n_2}}$ . Thus every  $n$ -variable Boolean function can be viewed as a mapping from  $\mathbb{F}_{2^{n_1}} \times \mathbb{F}_{2^{n_2}}$  to  $\mathbb{F}_2$ . By Lagrange interpolation, we can express  $f \in \mathbb{B}_n$  as

$$f(x, y) = \sum_{(a, b) \in \mathbb{F}_{2^{n_1}} \times \mathbb{F}_{2^{n_2}}} f(a, b) [1 + (x + a)^{2^{n_1}-1}]$$

$$\times [1 + (y + b)^{2^{n_2} - 1}]$$

To expand this expression, we should do operations (multiplications and additions) of elements from  $\mathbb{F}_{2^{n_1}}$  and  $\mathbb{F}_{2^{n_2}}$ . The smallest field in which these operations can be done is the composite field of  $\mathbb{F}_{2^{n_1}}$  and  $\mathbb{F}_{2^{n_2}}$ , i.e.  $\mathbb{F}_{2^{[n_1, n_2]}}$ , where “[ $\cdot$ ,  $\cdot$ ]” represents the least common multiple of two integers. Hence  $f$  can actually be represented into the form

$$f(x, y) = \sum_{i=0}^{2^{n_1}-1} \sum_{j=0}^{2^{n_2}-1} f_{i,j} x^i y^j, \quad f_{i,j} \in \mathbb{F}_{2^{[n_1, n_2]}}. \quad (1)$$

We call this kind of representation the bivariate representation of  $f$  over  $\mathbb{F}_{2^{n_1}} \times \mathbb{F}_{2^{n_2}}$ . It follows that as  $\mathbb{F}_2$ -algebras,  $\mathbb{B}_n$  can be viewed as a subalgebra of  $\mathbb{F}_{2^{[n_1, n_2]}}[x, y] / \langle x^{2^{n_1}} + x, y^{2^{n_2}} + y \rangle$ . Comparing dimensions we can also deduce the isomorphism

$$\mathbb{F}_{2^{[n_1, n_2]}}[x, y] / \langle x^{2^{n_1}} + x, y^{2^{n_2}} + y \rangle \cong \mathbb{B}_n \otimes_{\mathbb{F}_2} \mathbb{F}_{2^{[n_1, n_2]}}.$$

To obtain the ANF of  $f$  from its bivariate representation, we just need to choose two bases  $\{\alpha_1, \dots, \alpha_{n_1}\}$  and  $\{\beta_1, \dots, \beta_{n_2}\}$  of  $\mathbb{F}_{2^{n_1}}$  and  $\mathbb{F}_{2^{n_2}}$  over  $\mathbb{F}_2$  respectively, and write  $x = \sum_{i=1}^{n_1} x_i \alpha_i$ ,  $y = \sum_{j=1}^{n_2} y_j \beta_j$  for two sets of variables  $x_1, \dots, x_{n_1}$  and  $y_1, \dots, y_{n_2}$  over  $\mathbb{F}_2$ , and then put them into  $f(x, y)$ . It can be easily observed from this process that

$$\deg f \leq \max_{\substack{0 \leq i \leq 2^{n_1}-1 \\ 0 \leq j \leq 2^{n_2}-1}} \{\text{wt}_{n_1}(i) + \text{wt}_{n_2}(j) \mid f_{i,j} \neq 0\}.$$

The following lemma confirms that “=” actually holds.

**Proposition 1.** Assume  $n = n_1 + n_2$  and  $f \in \mathbb{B}_n$  with the bivariate representation (1). Then

$$\deg f = \max_{\substack{0 \leq i \leq 2^{n_1}-1 \\ 0 \leq j \leq 2^{n_2}-1}} \{\text{wt}_{n_1}(i) + \text{wt}_{n_2}(j) \mid f_{i,j} \neq 0\}.$$

*Proof:* Denote  $\mathbb{F}_{2^{[n_1, n_2]}}[x, y] / \langle x^{2^{n_1}} + x, y^{2^{n_2}} + y \rangle$  by  $\mathcal{R}_n$  and let  $\mathbb{R}_n$  be the  $\mathbb{F}_2$ -subalgebra of  $\mathcal{R}_n$  which is isomorphism to  $\mathbb{B}_n$ . For any  $0 \leq d \leq n$ , let  $R_d = \{h \in \mathbb{R}_n \mid h = \sum_{i,j} h_{i,j} x^i y^j, \text{wt}_{n_1}(i) + \text{wt}_{n_2}(j) \leq d \text{ for all } i, j \text{ with } h_{i,j} \neq 0\}$  and  $B_d = \{h \in \mathbb{B}_n \mid \deg h \leq d\}$ , which are  $\mathbb{F}_2$ -subspaces of  $\mathbb{R}_n$  and  $\mathbb{B}_n$  respectively. We just need to prove that  $\dim_{\mathbb{F}_2} R_d = \dim_{\mathbb{F}_2} B_d$  for all  $0 \leq d \leq n$ . First it is easy to see that

$$\dim_{\mathbb{F}_2} B_d = \sum_{k=0}^d \binom{n}{k} = \sum_{k=0}^d \binom{n_1 + n_2}{k}.$$

To get  $\dim_{\mathbb{F}_2} R_d$ , we note that  $\bar{R}_d = R_d \otimes_{\mathbb{F}_2} \mathbb{F}_{2^{[n_1, n_2]}}$  where

$$\bar{R}_d := \left\{ h \in \mathcal{R}_n \mid \begin{array}{l} h = \sum_{i,j} h_{i,j} x^i y^j, \\ \text{wt}_{n_1}(i) + \text{wt}_{n_2}(j) \leq d \\ \text{for all } i, j \text{ with } h_{i,j} \neq 0 \end{array} \right\}.$$

In fact, this can be observed from the isomorphism  $\mathbb{R}_n \otimes_{\mathbb{F}_2} \mathbb{F}_{2^{[n_1, n_2]}} = \mathcal{R}_n$  because essentially the “ $\otimes_{\mathbb{F}_2} \mathbb{F}_{2^{[n_1, n_2]}}$ ” operation only extends the definitional domain of coefficients of terms of functions in  $R_d$  to extend  $R_d$  to be an  $\mathbb{F}_{2^{[n_1, n_2]}}$ -vector space (more precisely, if  $R_d$  is spanned by a basis  $\{\beta_i\}$  over  $\mathbb{F}_2$ , then  $\bar{R}_d$  is spanned by the same basis over

$\mathbb{F}_{2^{[n_1, n_2]}}$ ), but all these terms  $(x^i y^j)$ ’s and the corresponding  $(\text{wt}_{n_1}(i) + \text{wt}_{n_2}(j))$ ’s are not affected. Therefore, we have

$$\dim_{\mathbb{F}_2} R_d = \dim_{\mathbb{F}_{2^{[n_1, n_2]}}} \bar{R}_d = \sum_{0 \leq k_1 + k_2 \leq d} \binom{n_1}{k_1} \binom{n_2}{k_2}.$$

By the Vandermonde’s convolution for binomial coefficients [7], we have

$$\sum_{0 \leq k \leq d} \binom{n_1 + n_2}{k} = \sum_{0 \leq k_1 + k_2 \leq d} \binom{n_1}{k_1} \binom{n_2}{k_2}.$$

This completes the proof.  $\blacksquare$

**Remark 1.** One may intuitively think the result of Proposition 1 natural. In fact, when  $n_1 = n_2 = n/2$  for an even integer  $n$ , the bivariate representations of Boolean functions in this case were frequently used in some authors’ work (see e.g. [15], [14], [8], [11]), and in all these work Proposition 1 was considered conventional and obvious, and was used without given a proof of it. However, we can see from the proof of Proposition 1 that, even for the above simple case, this result is far from obvious.

### C. Walsh transform of Boolean functions

The Walsh transform of a Boolean function is a useful tool in studying properties of it. The background of this concept is Fourier analysis on finite Abelian groups. In nature, for a Boolean function  $f$ , its Walsh transform is the Fourier transform of the complex valued function  $(-1)^f$  on a finite Abelian group. More precisely, for  $f \in \mathbb{B}_n$ , its Walsh transform at any  $\mathbf{a} \in \mathbb{F}_2^n$  can be defined as

$$W_f(\mathbf{a}) = \sum_{\mathbf{x} \in \mathbb{F}_2^n} (-1)^{f(\mathbf{x}) + \mathbf{a} \cdot \mathbf{x}} = \sum_{\mathbf{x} \in \mathbb{F}_2^n} (-1)^{f(\mathbf{x})} \chi_{\mathbf{a}}(\mathbf{x}),$$

where “ $\cdot$ ” represents the Euclidean inner product of vectors and  $\chi_{\mathbf{a}}$  is defined by  $\chi_{\mathbf{a}}(\mathbf{x}) = (-1)^{\mathbf{a} \cdot \mathbf{x}}$ ,  $\forall \mathbf{x} \in \mathbb{F}_2^n$ . This is because the dual group  $\mathbb{F}_2^n$  of the additive Abelian group  $\mathbb{F}_2^n$ , i.e. the group formed by all additive characters of  $\mathbb{F}_2^n$ , is actually  $\{\chi_{\mathbf{a}} \mid \mathbf{a} \in \mathbb{F}_2^n\}$ , all elements of which forms a standard orthogonal basis of the space formed by all functions from the group  $\mathbb{F}_2^n$  to  $\mathbb{C}^*$ , the multiplication group of the complex field. The Fourier transform of the complex valued function  $(-1)^f$  at  $\boldsymbol{\lambda} \in \mathbb{F}_2^n$  is in fact the coefficient before the term  $\chi_{\boldsymbol{\lambda}}$  of the Fourier expansion (i.e. the expansion under the basis  $\{\chi_{\mathbf{a}} \mid \mathbf{a} \in \mathbb{F}_2^n\}$ ) of  $(-1)^f$ . By this definition, it can be easily derived that  $f$  is balanced if and only if  $W_f(\mathbf{0}) = 0$ , and the nonlinearity of  $f$  can be equivalently expressed as

$$\mathcal{N}_f = 2^{n-1} - \frac{1}{2} \max_{\mathbf{a} \in \mathbb{F}_2^n} |W_f(\mathbf{a})|.$$

According to the meaning of Walsh transform explained above, we are clear that the Walsh transform of  $f \in \mathbb{B}_n$  at any  $\mathbf{a} \in \mathbb{F}_{2^n}$  can be defined as

$$W_f(\mathbf{a}) = \sum_{\mathbf{x} \in \mathbb{F}_{2^n}} (-1)^{f(\mathbf{x}) + \text{tr}_1^n(\mathbf{a}\mathbf{x})},$$

where  $\text{tr}_1^n(\cdot)$  is the trace function from  $\mathbb{F}_{2^n}$  to  $\mathbb{F}_2$ , i.e.  $\text{tr}_1^n(x) = \sum_{i=0}^{n-1} x^{2^i}$  for any  $x \in \mathbb{F}_{2^n}$ . This is because in this case the

dual group of  $\mathbb{F}_{2^n}$  is  $\widehat{\mathbb{F}_{2^n}} = \{\chi_a \mid a \in \mathbb{F}_{2^n}\}$  where for any  $a \in \mathbb{F}_{2^n}$ ,  $\chi_a(x) := (-1)^{\text{tr}_1^{n_1}(ax)}$ ,  $\forall x \in \mathbb{F}_{2^n}$ . Furthermore, when  $n = n_1 + n_2$  and  $f$  is viewed as a function from  $\mathbb{F}_{2^{n_1}} \times \mathbb{F}_{2^{n_2}}$  to  $\mathbb{F}_2$ , the Walsh transform of  $f$  at any  $(a, b) \in \mathbb{F}_{2^{n_1}} \times \mathbb{F}_{2^{n_2}}$  can be defined as

$$W_f(a, b) = \sum_{(x, y) \in \mathbb{F}_{2^{n_1}} \times \mathbb{F}_{2^{n_2}}} (-1)^{f(x, y) + \text{tr}_1^{n_1}(ax) + \text{tr}_1^{n_2}(by)}.$$

This is because in this case

$$(\mathbb{F}_{2^{n_1}} \times \mathbb{F}_{2^{n_2}})^\wedge = \widehat{\mathbb{F}_{2^{n_1}}} \times \widehat{\mathbb{F}_{2^{n_2}}} = \{\chi_a \cdot \psi_b \mid a \in \mathbb{F}_{2^{n_1}}, b \in \mathbb{F}_{2^{n_2}}\},$$

where for any  $a \in \mathbb{F}_{2^{n_1}}$ ,  $b \in \mathbb{F}_{2^{n_2}}$ ,  $\chi_a(x) := (-1)^{\text{tr}_1^{n_1}(ax)}$ ,  $\psi_b(y) := \text{tr}_1^{n_2}(by)$ ,  $\forall x \in \mathbb{F}_{2^{n_1}}$ ,  $y \in \mathbb{F}_{2^{n_2}}$ , according to the following lemma (see e.g. [9, Exercise 5.4]), the proof of which is simple and will be omitted.

**Lemma 1.** *Let  $G_1, G_2$  be two Abelian groups. Then  $\widehat{G_1 \times G_2} \cong \widehat{G_1} \times \widehat{G_2}$ .*

Similarly, we also have such equivalent expression of the nonlinearity of  $f$  as

$$\mathcal{N}_f = 2^{n-1} - \frac{1}{2} \max_{(a, b) \in \mathbb{F}_{2^{n_1}} \times \mathbb{F}_{2^{n_2}}} |W_f(a, b)|.$$

#### D. Algebraic immunity and immunity against FAA's of Boolean functions

The notion of algebraic immunity of Boolean functions was introduced in [13] to measure the ability of LFSR-based pseudo-random generators resisting algebraic attacks.

**Definition 1.** *Let  $f, g \in \mathbb{B}_n$ .  $g$  is called an annihilator of  $f$  if  $fg = 0$ . The algebraic immunity of  $f$ ,  $\text{AI}(f)$ , is defined to be the smallest possible degree of the nonzero annihilators of  $f$  or  $f + 1$ , i.e.*

$$\text{AI}(f) = \min_{0 \neq g \in \mathbb{B}_n} \{\deg(g) \mid fg = 0 \text{ or } (f + 1)g = 0\}.$$

It can be proved that the best possible value of the algebraic immunity of  $n$ -variable Boolean functions is  $\lceil n/2 \rceil$  [5], thus functions attaining this upper bound are often known as algebraic immunity optimal functions.

For a Boolean function  $f \in \mathbb{B}_n$ , optimal algebraic immunity is necessary but not sufficient since when there exists a function  $g$  of low degree such that  $gf$  is of a reasonable degree, a fast algebraic attack is feasible [6]. In fact,  $f$  is considered having best behavior against fast algebraic attacks if any pair of integers  $(e, d)$  with  $e < n/2$  and  $e + d < n$  such that there exists a nonzero function  $g$  of degree  $e$  satisfying that  $gf$  is of degree  $d$ , does not exist.

### III. GENERALIZED TU-DENG CONJECTURE

In [15] Tu and Deng proposed a combinatorial conjecture on binary strings (known as the Tu-Deng conjecture now), based on which they constructed a class of Boolean functions with optimal algebraic immunity.

**Conjecture 1 (Tu-Deng).** *Let  $n = 2k$  be an integer where  $k \geq 2$ . For any  $0 \leq t \leq 2^k - 2$ , define*

$$S_t = \left\{ (a, b) \mid \begin{array}{l} 0 \leq a, b \leq 2^k - 2, \\ a + b \equiv t \pmod{2^k - 1}, \\ \text{wt}_k(a) + \text{wt}_k(b) \leq k - 1 \end{array} \right\}.$$

*Then  $|S_t| \leq 2^{k-1}$ .*

As indicated in [14, Remark 2], this conjecture can be generalized by replacing  $a$  by  $ua$  for any fixed integer  $u$  with  $(u, 2^k - 1)$ , and particularly, for the case  $u = -2^l$  for some integer  $l \geq 0$ , a proof of this generalized conjecture can be achieved [4], [8]. Constructions of functions with optimal algebraic immunity based on this generalized conjecture were also obtained in [8].

In the sequel we assume  $n = (r + 1)m$  for an odd integer  $r \geq 1$  and an integer  $m \geq 3$ , and pick an integer  $u$  with  $(u, 2^m - 1)$ . We propose a new combinatorial conjecture on binary strings which is a more wide generalization of Conjecture 1.

**Conjecture 2.** *For any  $0 \leq t \leq 2^m - 2$ , define*

$$S_t = \left\{ (a, b) \mid \begin{array}{l} 0 \leq a \leq 2^{rm} - 2, 0 \leq b \leq 2^m - 1, \\ ua + b \equiv t \pmod{2^m - 1}, \\ \text{wt}_{rm}(a) + \text{wt}_m(b) \leq n/2 - 1 \end{array} \right\}.$$

*Then  $|S_t| \leq 2^{rm-1}$ .*

**Remark 2.** *It is easy to see that Conjecture 2 generalizes the conjecture proposed in [14, Remark 2] (see also [8, Conjecture 3.3]) and of course, Conjecture 1. Indeed, the conjecture in [14, Remark 2] can be viewed as the  $r = 1$  case of Conjecture 2 since in this case, the cardinality of  $S_t$  will not be affected if the restriction  $0 \leq b \leq 2^m - 1$  is replaced by  $0 \leq b \leq 2^m - 2$  for any  $0 \leq t \leq 2^m - 2$ . Therefore, when  $r = 1$  and  $u = -2^l$  for some integer  $l \geq 0$ , the conjecture is true according to [4].*

We have checked the conjecture by computer experiments for (1)  $r = 3, m = 3, 4, 5, 6, 7$ ; (2)  $r = 5, m = 3, 4$ ; and (3)  $r = 7, m = 3$ , for any  $u$  with  $(u, 2^m - 1) = 1$ , and for  $r = 3, m = 8$  for  $u = 1$ . Seeking a proof of this conjecture, even the Tu-Deng conjecture which is a very special case of it, is completely open. In addition, in the case  $r > 1$  and  $u = -2^l$  for some integer  $l \geq 0$ , it seems difficult to prove this conjecture though this can be done for  $r = 1$ .

### IV. A CLASS OF UNBALANCED FUNCTIONS

In the sequel, we fix a primitive element  $\alpha$  of  $\mathbb{F}_{2^{rm}}$  and set  $\beta = \alpha^{(2^{rm}-1)/(2^m-1)}$ , which is a primitive element of  $\mathbb{F}_{2^m}$ . For any integer  $0 \leq s \leq 2^{rm} - 2$ , we denote  $\Delta_s = \{\alpha^i \mid s \leq i \leq s + 2^{rm-1} - 1\}$ .

**Construction 1.** *Let  $0 \leq s \leq 2^{rm} - 2$  be an integer. Define an  $n$ -variable Boolean function  $f : \mathbb{F}_{2^{rm}} \times \mathbb{F}_{2^m} \rightarrow \mathbb{F}_2$  by setting*

$$\text{supp}(f) = \{(\gamma y^u, y) \mid y \in \mathbb{F}_{2^m}^*, \gamma \in \Delta_s\}.$$



**Remark 3.** It is easy to see that the bivariate representation of  $f$  over  $\mathbb{F}_{2^{rm}} \times \mathbb{F}_{2^m}$  can be written as

$$f(x, y) = g\left(\frac{x}{y^u}\right),$$

where  $g$  is an  $(rm)$ -variable Boolean function with  $\text{supp}(g) = \Delta_s$  (note that we always distinguish  $x/0$  with  $0$  in a finite field). We can see that this function can actually be viewed as a  $(2rm)$ -variable generalized Tu-Deng function (i.e. a function from [8, Construction 4.1]) with the second coordinate  $y$  limited to the subfield  $\mathbb{F}_{2^m}$  of  $\mathbb{F}_{2^{rm}}$ . In particular, when  $r = 1$ , it coincides with the unbalanced generalized Tu-Deng function (see [8, Construction 4.1]).

In the following we discuss some properties of the function defined in Construction 1.

#### A. Bivariate representation and algebraic degree

**Lemma 2.** Let  $h : \mathbb{F}_{2^{rm}} \times \mathbb{F}_{2^m} \rightarrow \mathbb{F}_2$  be an  $n$ -variable Boolean function. Then  $\deg h \leq n - 2$  if and only if  $\text{wt}(h)$  is even and

$$\sum_{(c_1, c_2) \in \text{supp}(h)} c_1 = \sum_{(c_1, c_2) \in \text{supp}(h)} c_2 = 0.$$

*Proof:* By Lagrange interpolation, the bivariate representation of  $h$  over  $\mathbb{F}_{2^{rm}} \times \mathbb{F}_{2^m}$  can be written as

$$\begin{aligned} h(x, y) &= \sum_{(c_1, c_2) \in \text{supp}(h)} \left[1 + (x + c_1)^{2^{rm}-1}\right] \left[1 + (x + c_2)^{2^m-1}\right] \\ &= |\text{supp}(h)| + \sum_{(c_1, c_2) \in \text{supp}(h)} (x + c_1)^{2^{rm}-1} \\ &\quad + \sum_{(c_1, c_2) \in \text{supp}(h)} (x + c_2)^{2^m-1} \\ &\quad + \sum_{(c_1, c_2) \in \text{supp}(h)} (x + c_1)^{2^{rm}-1} (x + c_2)^{2^m-1}. \end{aligned}$$

The coefficient of  $x^{2^{rm}-1}y^{2^m-1}$ , whose degree is  $n$ , is  $|\text{supp}(h)| \bmod 2$ ; the coefficients of  $x^{2^{rm}-2}y^{2^m-1}$  and  $x^{2^{rm}-1}y^{2^m-2}$ , whose degrees are  $n - 1$ , are  $\sum_{(c_1, c_2) \in \text{supp}(h)} c_1$  and  $\sum_{(c_1, c_2) \in \text{supp}(h)} c_2$  respectively. This completes the proof.  $\blacksquare$

**Lemma 3.** Let  $1 \leq j \leq (2^{rm}-1)/(2^m-1)-1$  be an integer. Then  $\text{wt}_{rm}((2^m-1)j) \leq rm - m$ .

*Proof:* For any  $1 \leq j \leq (2^{rm}-1)/(2^m-1)-1$ , it is obvious that

$$\text{wt}_{rm}\left((2^m-1)\left(\frac{2^{rm}-1}{2^m-1}-j\right)\right) = rm - \text{wt}_{rm}((2^m-1)j).$$

Thus we need only to prove that  $\text{wt}_{rm}((2^m-1)j) \geq m$  for any  $1 \leq j \leq (2^{rm}-1)/(2^m-1)-1$ . Without loss of generality, we can assume  $j$  is odd. Denote by  $\mathfrak{B}(a, b)$  the number of borrows when calculating  $a - b$  for two positive integers  $a$  and  $b$  with  $a \geq b$ . Then we have

$$\text{wt}_{rm}(2^m j - j) = \text{wt}_{rm}(2^m j) - \text{wt}_{rm}(j) + \mathfrak{B}(2^m j, j)$$

$$= \mathfrak{B}(2^m j, j).$$

It is easy to see that  $\mathfrak{B}(2^m j, j) \geq m$  since  $j$  is odd.  $\blacksquare$

**Theorem 1.** Let  $f$  be the Boolean function defined in Construction 1. Then the bivariate representation of  $f$  over  $\mathbb{F}_{2^{rm}} \times \mathbb{F}_{2^m}$  is

$$\begin{aligned} f(x, y) &= \sum_{\substack{i=1 \\ (2^m-1) \nmid i}}^{2^{rm}-2} \alpha^{-is} (1 + \alpha^{-i})^{2^{rm}-1} x^i y^{2^m-1-\overline{ui}} \\ &\quad + \sum_{j=1}^{\frac{2^{rm}-1}{2^m-1}-1} \alpha^{-(2^m-1)js} (1 + \alpha^{-(2^m-1)j})^{2^{rm}-1} \\ &\quad \times x^{(2^m-1)j} y^{2^m-1}, \end{aligned}$$

where  $\overline{ui}$  denotes the reduction of  $ui$  modulo  $(2^m - 1)$  in the residue class  $\{0, 1, \dots, 2^m - 2\}$  for any integer  $1 \leq i \leq 2^{rm} - 2$ . Therefore,  $n - m \leq \deg f \leq n - 2$ .

*Proof:* From the proof of [14, Theorem 2] we know that the univariate representation of the  $(rm)$ -variable function  $g$  defined in Remark 3 is

$$g(x) = \sum_{i=1}^{2^{rm}-2} \alpha^{-is} (1 + \alpha^{-i})^{2^{rm}-1} x^i.$$

Then the bivariate representation of  $f$  follows from Remark 3.

The algebraic degree of  $f$  is  $\max\{d_1, d_2\}$ , where

$$d_1 = \max \left\{ \text{wt}_{rm}(i) + \text{wt}_m(2^m - 1 - \overline{ui}) \mid \begin{array}{l} 1 \leq i \leq 2^{rm} - 2, \\ (2^m - 1) \nmid i \end{array} \right\}$$

and

$$d_2 = \max \left\{ \text{wt}_{rm}((2^m - 1)j) + m \mid 1 \leq j \leq \frac{2^{rm}-1}{2^m-1} - 1 \right\}.$$

By Lemma 2 we can get  $d_1, d_2 \leq n - 2$ . When  $i = 2^{rm} - 2$ ,  $\text{wt}_{rm}(i) + \text{wt}_m(2^m - 1 - \overline{ui}) = rm - 1 + m - \text{wt}_m(\overline{ui}) = n - (\text{wt}_m(\overline{u}) + 1)$ , hence  $n - m \leq d_1 \leq n - 2$ . On the other hand, when  $j = (2^{rm}-1)/(2^m-1)-1$ ,  $\text{wt}_{rm}((2^m-1)j) = rm - m$ , hence we have  $d_2 = rm = n - m$  from Lemma 3. Finally we get that  $n - m \leq \deg f \leq n - 2$ .  $\blacksquare$

**Remark 4.** From the proof of Theorem 1 we can see that:

- (1) when  $u = 2^t$  for some non-negative integer  $t$ ,  $\deg f = n - m$ ; and
- (2) when  $u = -2^t$  for some non-negative integer  $t$ ,  $\deg f = n - 2$ .

**Corollary 1.** Let  $f$  be the Boolean function defined in Construction 1. Then  $f$  is bent if and only if  $r = 1$  and  $u = 2^t$  for some non-negative integer  $t$ .

*Proof:* Since the algebraic degree of an  $n$ -variable bent function is at most  $n/2$  and  $n/2 \leq n - m \leq \deg f \leq n - 2$  from Theorem 1, we know that only when  $r = 1$ , i.e.  $n - m = n/2$ ,  $f$  is possibly bent. Furthermore, when  $u = 2^t$  for some non-negative integer  $t$ , it is clear that  $f$  is bent (in fact,  $f$  is equivalent to a function belonging to the well-known

$\mathcal{PS}_{\text{ap}}$  class of bent functions). To prove this condition is also necessary, we should prove that  $\deg f = n/2 = m$  implies  $\text{wt}_m(u) = 1$ . In fact, for any  $1 \leq i \leq 2^m - 2$ ,  $\text{wt}_m(i) + \text{wt}_m(2^m - 1 - \overline{ui}) = m + \text{wt}_m(i) - \text{wt}_m(\overline{ui}) \leq \deg f = m$ , thus we have  $\text{wt}_m(i) \leq \text{wt}_m(\overline{ui})$ . Fixing  $i$  to be  $2^m - 2$ , we get  $m - 1 \leq \text{wt}_m(\overline{u}) = m - \text{wt}_m(u)$ , which implies that  $\text{wt}_m(u) \leq 1$ . Therefore,  $\text{wt}_m(u) = 1$ . ■

### B. Algebraic immunity

**Theorem 2.** *Let  $f$  be the Boolean function defined in Construction 1. Then  $\text{AI}(f) \leq m$ . In particular,  $f$  has optimal algebraic immunity provided that Conjecture 2 is true when  $r = 1$ .*

*Proof:* Obviously,  $1 + y^{2^m-1}$  is an annihilator of  $f$ , whose degree is  $m$ . This implies that  $\text{AI}(f) \leq m$ .

When  $r = 1$ ,  $f$  coincides with the function defined in [8, Construction 4.1] and Conjecture 2 coincides with [8, Conjecture 3.3] according to Remark 3 and Remark 2 respectively, so from [8, Theorem 4.2] we are clear that  $\text{AI}(f) = m = n/2$  if Conjecture 2 is true. ■

From Theorem 2 we can see that the algebraic immunity of the functions from Construction 1 is not possible to be optimal when  $r > 1$ . However, it is interesting that they can be modified to be functions with optimal algebraic immunity when modified to be balanced functions. So in this case, our process to obtain balanced functions with optimal algebraic immunity is different from those in [15], [14], [8], where balanced functions with optimal algebraic immunity were all modified from unbalanced ones with optimal algebraic immunity.

## V. A CLASS OF BALANCED FUNCTIONS WITH GOOD CRYPTOGRAPHIC PROPERTIES

**Construction 2.** *Let  $0 \leq s, l \leq 2^{rm} - 2$  be two integers. Define an  $n$ -variable Boolean function  $F : \mathbb{F}_{2^{rm}} \times \mathbb{F}_{2^m} \rightarrow \mathbb{F}_2$  by setting*

$$\begin{aligned} \text{supp}(F) = & \{(\gamma y^u, y) \mid y \in \mathbb{F}_{2^m}^*, \gamma \in \Delta_s\} \\ & \cup \{(\gamma, 0) \mid \gamma \in \Delta_l\}. \end{aligned}$$

**Remark 5.** *It is easy to see that the bivariate representation of  $F$  over  $\mathbb{F}_{2^{rm}} \times \mathbb{F}_{2^m}$  can be written as*

$$F(x, y) = \begin{cases} g\left(\frac{x}{y^u}\right) & \text{if } xy \neq 0 \\ \omega(x) & \text{if } y = 0, \end{cases}$$

where  $g$  and  $\omega$  are  $(rm)$ -variable functions with  $\text{supp}(g) = \Delta_s$  and  $\text{supp}(\omega) = \Delta_l$ .

**Remark 6.** *It is easy to see that if  $u_1$  and  $u_2$  are chosen from the same cyclotomic coset modulo  $(2^m - 1)$ , then the functions defined from  $u_1$  and  $u_2$  in Construction 2 are linearly equivalent.*

Note that Construction 2 provides various ways to obtain  $n$ -variable Boolean functions for an even integer  $n$  since the parameters, namely  $m, u, s$  and  $l$ , can be flexibly chosen. In the following, we discuss some cryptographic properties of the function  $F$ .

### A. Balancedness, bivariate representation and algebraic degree

**Theorem 3.** *Let  $F$  be the Boolean function defined in Construction 2. Then  $F$  is balanced.*

*Proof:* It is obvious that  $|\text{supp}(f)| = (2^m - 1)2^{rm-1} + 2^{rm-1} = 2^{n-1}$ , so  $F$  is balanced. ■

**Theorem 4.** *Let  $F$  be the Boolean function defined in Construction 2. Then the bivariate representation of  $F$  over  $\mathbb{F}_{2^{rm}} \times \mathbb{F}_{2^m}$  is*

$$\begin{aligned} F(x, y) = & \sum_{\substack{i=1 \\ (2^m-1) \nmid i}}^{2^{rm}-2} \alpha^{-is} (1 + \alpha^{-i})^{2^{rm-1}-1} x^i y^{2^m-1-\overline{ui}} \\ & + \sum_{j=1}^{\frac{2^{rm}-1}{2^m}-1} \alpha^{-(2^m-1)js} (1 + \alpha^{-(2^m-1)j})^{2^{rm-1}-1} \\ & \quad \times x^{(2^m-1)j} y^{2^m-1} \\ & + \sum_{i=1}^{2^{rm}-2} \alpha^{-il} (1 + \alpha^{-i})^{2^{rm-1}-1} x^i (1 + y^{2^m-1}). \end{aligned}$$

Therefore,  $\deg F = n - 1$ , i.e.  $F$  has optimal algebraic degree.

*Proof:* It is easy to see from Remark 5 that the bivariate representation of  $F$  over  $\mathbb{F}_{2^{rm}} \times \mathbb{F}_{2^m}$  can be written as

$$F(x, y) = f(x, y) + \omega(x)(1 + y^{2^m-1}),$$

where  $f(x, y)$  is the function defined in Construction 1. Then the representation of  $F$  follows from Theorem 1 and [14, Theorem 2].

Since  $\omega(x)$  is in fact an  $(rm)$ -variable Carlet-Feng function, we are clear that  $\deg \omega = rm - 1$  according to [3, Theorem 2], so the degree of  $\omega(x)(1 + y^{2^m-1})$  is  $rm - 1 + m = n - 1$ . However, by Theorem 1 we have  $\deg f \leq n - 2$ . Finally we know that  $\deg F = n - 1$ , which is optimal for a balanced function. ■

### B. Algebraic immunity

In this subsection, we study the algebraic immunity of the functions from Construction 2. For the basic notions about BCH codes and related results that will be used in the proof, we refer to [12]. Besides, the following lemma is also necessary.

**Lemma 4.** *Let  $U, V \in \mathbb{F}_2^t$  be two binary vectors. Then  $\text{wt}(U) + \text{wt}(V) \geq \text{wt}(U + V)$ .*

*Proof:* It is easy to see that  $\text{wt}(U + V) = \text{wt}(U) + \text{wt}(V) - \text{wt}(U \times V)$ , where  $U \times V$  represents the Hadamard product (i.e. bitwise multiplication) of  $U$  and  $V$ . ■

**Theorem 5.** *Let  $F$  be the Boolean function defined in Construction 2. Then  $F$  has optimal algebraic immunity provided that Conjecture 2 is true.*

*Proof:* Since when  $r = 1$  the proof is almost the same with the proof of [8, Theorem 5.3], we need only to deal with the case  $r > 1$ . We proceed by proving both  $F$  and  $F + 1$  have

no nonzero annihilators of degree less than  $n/2$  if Conjecture 2 is true.

Assume  $h$  is an  $n$ -variable Boolean function with  $\deg h < n/2$  and  $hF = 0$ . Write  $h$  into its bivariate representation over  $\mathbb{F}_{2^{rm}} \times \mathbb{F}_{2^m}$  as

$$h(x, y) = \sum_{i=0}^{2^{rm}-1} \sum_{j=0}^{2^m-1} h_{i,j} x^i y^j.$$

From  $\deg h < n/2 < rm$  we know that  $h_{i,j} = 0$  for any  $i, j$  with  $\text{wt}_{rm}(i) + \text{wt}_m(j) \geq n/2$ , which implies  $h_{2^{rm}-1,j} = 0$  for any  $0 \leq j \leq 2^m - 1$ . Thus we can write  $h$  into the form

$$h(x, y) = \sum_{i=0}^{2^{rm}-2} \sum_{j=0}^{2^m-2} h_{i,j} x^i y^j + \sum_{i=0}^{2^{rm}-2} h_{i,2^m-1} x^i y^{2^m-1}.$$

From  $h|_{\text{supp}(F)} = 0$  we get that, for any  $y \in \mathbb{F}_{2^m}^*, \gamma \in \Delta_s$ ,

$$\begin{aligned} h(\gamma y^u, y) &= \sum_{i=0}^{2^{rm}-2} \sum_{j=0}^{2^m-2} h_{i,j} \gamma^i y^{ui+j} + \sum_{i=0}^{2^{rm}-2} h_{i,2^m-1} \gamma^i y^{ui} \\ &= \sum_{k=0}^{2^m-2} y^k \left[ \sum_{i=0}^{2^{rm}-2} h_{i,k-ui \pmod{2^m-1}} \gamma^i \right. \\ &\quad \left. + \sum_{j=0}^{\frac{2^{rm}-1}{2^m}-1} h_{\tilde{u}k+j(2^m-1), 2^m-1} \gamma^{\tilde{u}k+j(2^m-1)} \right] \\ &= \sum_{k=0}^{2^m-2} h_k(\gamma) y^k \\ &= 0, \end{aligned}$$

where  $\tilde{u}$  is the integer satisfying  $u\tilde{u} \equiv 1 \pmod{2^m-1}$  and  $0 \leq \tilde{u}k \leq 2^m - 2$  is considered modulo  $(2^m - 1)$ , and

$$\begin{aligned} h_k(\gamma) &= \sum_{i=0}^{2^{rm}-2} h_{i,k-ui \pmod{2^m-1}} \gamma^i \\ &\quad + \sum_{j=0}^{\frac{2^{rm}-1}{2^m}-1} h_{\tilde{u}k+j(2^m-1), 2^m-1} \gamma^{\tilde{u}k+j(2^m-1)}. \end{aligned}$$

Therefore, for any  $0 \leq k \leq 2^m - 2$ ,  $h_k(\gamma) = 0$  for any  $\gamma \in \Delta_s$ . Viewing  $h_k(\gamma)$  as a polynomial in  $\gamma$ , we find that the vector of coefficients can be represented as

$$\begin{aligned} \mathbf{h}_k &= (h_{0,k}, h_{1,k-u}, \dots, h_{\tilde{u}k,0}, \dots, h_{2^m-2,k+u}, \\ &\quad h_{2^m-1,k}, h_{2^m,k-u}, \dots, h_{2^m-1+\tilde{u}k,0}, \dots, h_{2^m+1-3,k+u}, \\ &\quad \dots, h_{2^{rm}-2m+\tilde{u}k,0}, \dots, h_{2^{rm}-2,k+u}) \\ &\quad + (0, \dots, 0, h_{\tilde{u}k,2^m-1}, 0, \dots, 0, h_{2^m-1+\tilde{u}k,2^m-1}, 0, \\ &\quad \dots, 0, h_{2^{rm}-2m+\tilde{u}k,2^m-1}, 0, \dots, 0) \\ &:= \mathbf{h}_k^{(1)} + \mathbf{h}_k^{(2)}. \end{aligned}$$

Now that  $\mathbf{h}_k$  can be viewed as a codeword of certain BCH code with designed distance  $2^{rm-1} + 1$ , if it is not zero, the BCH bound implies that  $\text{wt}(\mathbf{h}_k) \geq 2^{rm-1} + 1$ . On the other hand, Lemma 4 and Conjecture 2 imply that

$$\text{wt}(\mathbf{h}_k) = \text{wt}(\mathbf{h}_k^{(1)} + \mathbf{h}_k^{(2)}) \leq \text{wt}(\mathbf{h}_k^{(1)}) + \text{wt}(\mathbf{h}_k^{(2)}) \leq 2^{rm-1}.$$

A contradiction follows and hence we have  $\mathbf{h}_k = 0$  for any  $0 \leq k \leq 2^m - 2$ , which leads to the fact that  $h_{i,0} = h_{i,2^m-1}$  for any  $0 \leq i \leq 2^{rm} - 2$  with  $i \equiv \tilde{u}k \pmod{2^m-1}$ , and  $h_{i,k-\tilde{u}i} = 0$  otherwise. Since we have the equality

$$\begin{aligned} \bigcup_{0 \leq k \leq 2^m-2} \{0 \leq i \leq 2^{rm}-2 \mid i \equiv \tilde{u}k \pmod{2^m-1}\} \\ = \{i \mid 0 \leq i \leq 2^{rm}-2\}, \end{aligned}$$

we are now clear that the annihilator  $h$  is of the form

$$\begin{aligned} h(x, y) &= \sum_{i=0}^{2^{rm}-2} (h_{i,0} x^i + h_{i,2^m-1} x^i y^{2^m-1}) \\ &= (1 + y^{2^m-1}) \sum_{i=0}^{2^{rm}-2} h_{i,0} x^i. \end{aligned}$$

In fact, the sums above are over all  $i$ 's with  $\text{wt}_{rm}(i) < n/2 - m$ . Noting that  $\{(\gamma, 0) \mid \gamma \in \Delta_l\} \subseteq \text{supp}(f)$ , we have  $h(\gamma, 0) = 0$  for any  $\gamma \in \Delta_l$ , that is

$$\sum_{i=0}^{2^{rm}-2} h_{i,0} \gamma^i = 0 \text{ for any } \gamma \in \Delta_l.$$

Denote  $\mathbf{h}' = (h_{0,0}, h_{1,0}, \dots, h_{2^{rm}-2,0})$ . If  $\mathbf{h}' \neq \mathbf{0}$ , the BCH bound implies that  $\text{wt}(\mathbf{h}') \geq 2^{rm-1} + 1$ ; on the other hand, the restriction on the degree of  $h$  leads to

$$\text{wt}(\mathbf{h}') \leq \sum_{k=0}^{n/2-m-1} \binom{rm}{k} < \sum_{k=0}^{\lfloor \frac{rm-1}{2} \rfloor} \binom{rm}{k} \leq 2^{rm-1}.$$

This contradiction implies that  $\mathbf{h}' = \mathbf{0}$ , i.e.  $h = 0$ .

As for  $f + 1$ , the proof is almost the same. Assume  $h$  is a Boolean function with  $\deg h < n/2$  and  $h(f + 1) = 0$  represented as above. In this case,  $h(\gamma y^u, y) = 0$  for any  $\gamma \in \mathbb{F}_{2^{rm}}^* \setminus \Delta_s, y \in \mathbb{F}_{2^m}^*$ , thus  $\mathbf{h}_k$  can be viewed as a codeword of certain BCH code with designed distance  $2^{rm-1}$  and the BCH bound implies that  $\text{wt}(\mathbf{h}_k) \geq 2^{rm-1}$  if  $\mathbf{h}_k \neq \mathbf{0}$ . On the other hand,  $h(0, y) = 0$  for any  $y \in \mathbb{F}_{2^m}^*$ , which implies that  $h_{0,k} = 0$  for all  $0 \leq k \leq 2^m - 1$ . Since  $\text{wt}_m(k) \leq m - 1 < n/2 - 1$  for any  $0 \leq k \leq 2^m - 2$ , Lemma 4 together with Conjecture 2 imply that  $\text{wt}(\mathbf{h}_k) \leq 2^{rm-1} - 1$ , which lead to a contradiction. Then we get that  $h$  is of the form

$$h(x, y) = (1 + y^{2^m-1}) \sum_{i=0}^{2^{rm}-2} h_{i,0} x^i.$$

Further noting that  $h(\gamma, 0) = 0$  for any  $\gamma \in \mathbb{F}_{2^{rm}}^* \setminus \Delta_l$ , we get  $\text{wt}(\mathbf{h}') \geq 2^{rm-1}$  by the BCH bound if  $\mathbf{h}' \neq \mathbf{0}$ , where  $\mathbf{h}' = (h_{0,0}, h_{1,0}, \dots, h_{2^{rm}-2,0})$ . However, from the restriction on the degree of  $h$ , we have  $\text{wt}(\mathbf{h}') \leq 2^{rm-1} - 1$ . This contradiction leads to  $h = 0$ . We complete the proof. ■

**Remark 7.** Set  $D_{rm} = \sum_{k=0}^{n/2-m-1} \binom{rm}{k}$  and  $\Theta_t = \{\alpha^t, \dots, \alpha^{t+D_{rm}-1}\}$  for any integer  $0 \leq t \leq 2^{rm} - 2$ . Assume  $l$  and  $l'$  satisfy that  $\Theta_l \cap \Theta_{l'} = \emptyset$ . Then from the proof of Theorem 5, it can be observed that if we set  $\text{supp}(\omega) = \{(\gamma, 0) \mid \gamma \in \Theta_l \cup C\}$  where  $C$  is any subset of  $\mathbb{F}_{2^{rm}}^* \setminus (\Theta_l \cup \Theta_{l'})$  with size  $2^{rm-1} - D_{rm}$ , the function  $F$

constructed with this  $\omega$  will also be balanced and have optimal algebraic immunity provided Conjecture 2 is true. However, the algebraic degree of functions constructed in this manner might not be optimal.

### C. Nonlinearity

Applying the classical technique of using Gauss sums to estimate nonlinearity of Boolean functions constructed based on finite fields (see, for example, [3], [15], [8] and especially [14], [11]), we can also obtain a lower bound of the nonlinearity of the functions from Construction 2. For simplicity, we use "Tr" and "tr" to denote "tr<sub>1<sup>m</sup></sub>" and "tr<sub>1</sub>" respectively and denote  $Q = 2^{rm}$ ,  $q = 2^m$ .

**Lemma 5** ([14]). *For every  $0 < x < \pi/2$ ,*

$$\frac{1}{x} + \frac{x}{6} < \frac{1}{\sin x} < \frac{1}{x} + \frac{x}{4}.$$

**Lemma 6.** *Let  $T \geq 2$  be an integer. Then*

$$2T \left( \frac{\ln T}{\pi} + 0.163 \right) < \sum_{i=1}^{T-1} \frac{1}{\sin \frac{\pi i}{2T}} < 2T \left( \frac{\ln T}{\pi} + 0.263 \right) + \frac{3\pi}{8T}.$$

*Proof:* From Lemma 5 we have

$$\begin{aligned} \sum_{i=1}^{T-1} \frac{1}{\sin \frac{\pi i}{2T}} &> \frac{2T}{\pi} \sum_{i=1}^{T-1} \frac{1}{i} + \frac{\pi}{12T} \sum_{i=1}^{T-1} i \\ &\geq \frac{2T}{\pi} \left( 1 + \sum_{i=2}^{T-1} \int_i^{i+1} \frac{dz}{z} \right) + \frac{\pi(T-1)}{24} \\ &= \frac{2T}{\pi} \left( 1 + \int_2^T \frac{dz}{z} \right) + \frac{\pi(T-1)}{24} \\ &= \frac{2T}{\pi} (\ln T + 1 - \ln 2) + \frac{\pi(T-1)}{24} \\ &= 2T \left( \frac{\ln T}{\pi} + \frac{1 - \ln 2}{\pi} + \frac{\pi}{48} \right) - \frac{\pi}{24} \\ &> 2T \left( \frac{\ln T}{\pi} + 0.163 \right). \end{aligned}$$

On the other hand, we have

$$\begin{aligned} \sum_{i=1}^{T-1} \frac{1}{\sin \frac{\pi i}{2T}} &< \left( \frac{2T}{\pi} + \frac{\pi}{8T} + \frac{T}{\pi} + \frac{\pi}{4T} \right) + \frac{2T}{\pi} \sum_{i=3}^{T-1} \frac{\frac{\pi}{2T}}{\sin \frac{\pi i}{2T}} \\ &< \frac{3T}{\pi} + \frac{3\pi}{8T} + \frac{2T}{\pi} \sum_{i=3}^{T-1} \int_{\frac{\pi i}{2T} - \frac{\pi}{4T}}^{\frac{\pi i}{2T} + \frac{\pi}{4T}} \frac{dz}{\sin z} \\ &< \frac{3T}{\pi} + \frac{3\pi}{8T} + \frac{2T}{\pi} \int_{\frac{5\pi}{4T}}^{\frac{\pi}{2}} \frac{dz}{\sin z} \\ &= \frac{3T}{\pi} + \frac{3\pi}{8T} - \frac{2T}{\pi} \ln \left( \tan \frac{5\pi}{8T} \right) \\ &\leq \frac{3\pi}{8T} + \frac{2T}{\pi} \left( \ln T + 1.5 - \ln \frac{5\pi}{8T} \right) \\ &< 2T \left( \frac{\ln T}{\pi} + 0.263 \right) + \frac{3\pi}{8T}. \end{aligned}$$

**Lemma 7.** *Let  $0 \leq s \leq Q-2$  be an integer and*

$$\Lambda_s = \sum_{\gamma \in \Delta_s} \sum_{y \in \mathbb{F}_q^*} (-1)^{\text{Tr}(\gamma y)}.$$

*Then  $|\Lambda_s| = 2^{m-1}$  when  $r = 1$  and*

$$|\Lambda_s| \leq \left\lceil \frac{(n-2m) \ln 2}{\pi} + 0.263 \right\rceil 2^{(n-m)/2} + 2^{m-1} + 1$$

*when  $r > 1$ .*

*Proof:* Let  $\xi \in \mathbb{C}$  be a  $(Q-1)$ -th root of unity and  $\zeta = \xi^N$  where  $N = (Q-1)/(q-1)$ . Denote by  $\chi_1$  the primitive multiplication character of  $\mathbb{F}_Q^*$  and define the Gauss sums over  $\mathbb{F}_Q$  as

$$G_1(\chi_1^\mu) = \sum_{x \in \mathbb{F}_Q^*} \chi_1^\mu(x) (-1)^{\text{Tr}(x)}$$

for any  $0 \leq \mu \leq Q-2$ . It is well known that  $G_1(\chi_1^0) = -1$  and  $|G_1(\chi_1^\mu)| = Q^{1/2}$  for any  $1 \leq \mu \leq Q-2$  [9]. By Fourier inversion we have, for any  $0 \leq i \leq Q-2$ ,

$$(-1)^{\text{Tr}(\alpha^i)} = \frac{1}{Q-1} \sum_{\mu=0}^{Q-2} G_1(\chi_1^\mu) \xi^{-\mu i}.$$

Hence we have

$$\begin{aligned} \Lambda_s &= \sum_{i=s}^{s+\frac{Q}{2}-1} \sum_{j=0}^{q-2} (-1)^{\text{Tr}(\alpha^{i+Nj})} \\ &= \frac{1}{Q-1} \sum_{i=s}^{s+\frac{Q}{2}-1} \sum_{j=0}^{q-2} \sum_{\mu=0}^{Q-2} G_1(\chi_1^\mu) \xi^{-\mu(i+Nj)} \\ &= \frac{1}{Q-1} \sum_{\mu=0}^{Q-2} G_1(\chi_1^\mu) \sum_{i=s}^{s+\frac{Q}{2}-1} \xi^{-\mu i} \sum_{j=0}^{q-2} \zeta^{-\mu j}. \end{aligned}$$

Note that

$$\sum_{i=s}^{s+\frac{Q}{2}-1} \xi^{-\mu i} = \begin{cases} \frac{Q}{2} & \text{if } \mu = 0 \\ \xi^{-\mu s} \frac{1 - \xi^{-\mu \frac{Q}{2}}}{1 - \xi^{-\mu}} & \text{otherwise,} \end{cases}$$

$$\sum_{j=0}^{q-2} \zeta^{-\mu j} = \begin{cases} q-1 & \text{if } \mu \equiv 0 \pmod{q-1} \\ 0 & \text{otherwise.} \end{cases}$$

Then we have

$$\begin{aligned} \Lambda_s &= \frac{1}{Q-1} \left[ -\frac{Q(q-1)}{2} \right. \\ &\quad \left. + (q-1) \sum_{\substack{\mu=1 \\ (q-1) \mid \mu}}^{Q-2} G_1(\chi_1^\mu) \xi^{-\mu s} \frac{1 - \xi^{-\mu \frac{Q}{2}}}{1 - \xi^{-\mu}} \right]. \end{aligned}$$

Note that when  $r = 1$ , i.e.  $Q = q$ , the above formula yields  $\Lambda_s = -Q/2 = -2^{m-1}$ . When  $r > 1$ , we can get that

$$|\Lambda_s| \leq \frac{Q(q-1)}{2(Q-1)} + \frac{Q^{1/2}(q-1)}{Q-1} \sum_{\substack{\mu=1 \\ (q-1) \mid \mu}}^{Q-2} \left| \frac{1 - \xi^{-\mu \frac{Q}{2}}}{1 - \xi^{-\mu}} \right|$$

■



$$\begin{aligned}
&= \frac{Q(q-1)}{2(Q-1)} + \frac{Q^{1/2}(q-1)}{Q-1} \sum_{\substack{\mu=1 \\ (q-1) \nmid \mu}}^{Q-2} \left| \frac{1}{1 + \xi^{-\mu/2}} \right| \\
&= \frac{Q(q-1)}{2(Q-1)} + \frac{Q^{1/2}(q-1)}{Q-1} \sum_{k=1}^{N-1} \frac{1}{2 \sin \frac{\pi k}{2N}}.
\end{aligned}$$

By Lemma 6 we have

$$\begin{aligned}
&|\Lambda_s| \\
&\leq \frac{Q(q-1)}{2(Q-1)} + \frac{Q^{1/2}(q-1)}{2(Q-1)} \left[ 2N \left( \frac{\ln N}{\pi} + 0.263 \right) + \frac{3\pi}{8N} \right] \\
&< \frac{q}{2} + Q^{1/2} \left( \frac{1}{\pi} \ln \frac{Q}{q} + 0.263 \right) + \frac{3\pi Q^{1/2}(q-1)^2}{16(Q-1)^2} \\
&\leq \left[ \frac{(n-2m) \ln 2}{\pi} + 0.263 \right] 2^{(n-m)/2} + 2^{m-1} + 1.
\end{aligned}$$

■

**Lemma 8.** Let  $0 \leq s \leq Q-2$  be an integer. Denote

$$\Gamma_s = \sum_{\gamma \in \Delta_s} \sum_{y \in \mathbb{F}_q^*} (-1)^{\text{Tr}(\gamma y) + \text{tr}(y^u)},$$

where  $u$  is an integer with  $(u, q-1) = 1$ . Then

$$\begin{aligned}
|\Gamma_s| &\leq \left[ \frac{(n-m) \ln 2}{\pi} + 0.263 \right] 2^{n/2} \\
&\quad - \left[ \frac{(n-2m) \ln 2}{\pi} + 0.163 \right] 2^{n/2-m} + 2.
\end{aligned}$$

*Proof:* Notations the same as those in the proof of Lemma 7 and further assume  $\chi_2$  to be the primitive multiplication character of  $\mathbb{F}_q^*$ , and denote the Gauss sums over  $\mathbb{F}_q$  by  $G_2(\chi_2^\nu)$  for any  $0 \leq \nu \leq q-2$ , i.e.

$$G_2(\chi_2^\nu) = \sum_{x \in \mathbb{F}_q^*} \chi_2^\nu(x) (-1)^{\text{tr}(x)}.$$

We also have  $G_2(\chi_2^0) = -1$ ,  $|G_2(\chi_2^\mu)| = q^{1/2}$  for any  $1 \leq \nu \leq q-2$  and

$$(-1)^{\text{tr}(\beta^j)} = \frac{1}{q-1} \sum_{\nu=0}^{q-2} G_2(\chi_2^\nu) \zeta^{-\nu j}$$

for any  $0 \leq j \leq q-2$ . Hence we have

$$\begin{aligned}
\Gamma_s &= \sum_{i=s}^{s+\frac{Q}{2}-1} \sum_{j=0}^{q-2} (-1)^{\text{Tr}(\alpha^i \beta^j) + \text{tr}(\beta^j u)} \\
&= \frac{1}{(Q-1)(q-1)} \sum_{i=s}^{s+\frac{Q}{2}-1} \sum_{j=0}^{q-2} \sum_{\mu=0}^{Q-2} G_1(\chi_1^\mu) \xi^{-\mu(i+Nj)} \\
&\quad \times \sum_{\nu=0}^{q-2} G_2(\chi_2^\nu) \zeta^{-\nu j u} \\
&= \frac{1}{(Q-1)(q-1)} \sum_{\mu=0}^{Q-2} \sum_{\nu=0}^{q-2} G_1(\chi_1^\mu) G_2(\chi_2^\nu) \\
&\quad \times \sum_{i=s}^{s+\frac{Q}{2}-1} \xi^{-\mu i} \sum_{j=0}^{q-2} \zeta^{-(\nu u + \mu)j}.
\end{aligned}$$

Note that

$$\sum_{i=s}^{s+\frac{Q}{2}-1} \xi^{-\mu i} = \begin{cases} \frac{Q}{2} & \text{if } \mu = 0 \\ \xi^{-\mu s} \frac{1 - \xi^{-\mu \frac{Q}{2}}}{1 - \xi^{-\mu}} & \text{otherwise,} \end{cases}$$

$$\sum_{j=0}^{q-2} \zeta^{-(\nu u + \mu)j} = \begin{cases} q-1 & \text{if } \nu u + \mu \equiv 0 \pmod{q-1} \\ 0 & \text{otherwise.} \end{cases}$$

Since  $\nu u + \mu \equiv 0 \pmod{q-1}$  if and only if  $\nu = 0$  and  $\mu = k(q-1)$  for some  $0 \leq k \leq N-1$ , or  $\nu \equiv q-1-\tilde{u}\mu \pmod{q-1}$  and  $(q-1) \nmid \mu$  where  $\tilde{u}u \equiv 1 \pmod{q-1}$ , we have

$$\begin{aligned}
\Gamma_s &= \frac{1}{(Q-1)(q-1)} \left[ \frac{Q(q-1)}{2} \right. \\
&\quad + (q-1) \sum_{\substack{\mu=1 \\ (q-1) \nmid \mu}}^{Q-2} G_1(\chi_1^\mu) G_2(\chi_2^{q-1-\tilde{u}\mu}) \xi^{-\mu s} \frac{1 - \xi^{-\mu \frac{Q}{2}}}{1 - \xi^{-\mu}} \\
&\quad \left. + (q-1) \sum_{\substack{\mu=1 \\ (q-1) \mid \mu}}^{Q-2} G_1(\chi_1^\mu) (-1) \xi^{-\mu s} \frac{1 - \xi^{-\mu \frac{Q}{2}}}{1 - \xi^{-\mu}} \right].
\end{aligned}$$

Therefore, we can get that

$$\begin{aligned}
|\Gamma_s| &\leq \frac{Q}{2(Q-1)} + \frac{Q^{1/2}q^{1/2}}{Q-1} \sum_{\substack{\mu=1 \\ (q-1) \nmid \mu}}^{Q-2} \left| \frac{1 - \xi^{-\mu \frac{Q}{2}}}{1 - \xi^{-\mu}} \right| \\
&\quad + \frac{Q^{1/2}}{Q-1} \sum_{\substack{\mu=1 \\ (q-1) \mid \mu}}^{Q-2} \left| \frac{1 - \xi^{-\mu \frac{Q}{2}}}{1 - \xi^{-\mu}} \right| \\
&< 1 + \frac{Q^{1/2}q^{1/2}}{Q-1} \sum_{\mu=1}^{Q-2} \left| \frac{1}{1 + \xi^{-\mu/2}} \right| \\
&\quad - \frac{Q^{1/2}(q^{1/2}-1)}{Q-1} \sum_{\substack{\mu=1 \\ (q-1) \mid \mu}}^{Q-2} \left| \frac{1}{1 + \xi^{-\mu/2}} \right| \\
&\leq 1 + \frac{Q^{1/2}q^{1/2}}{Q-1} \sum_{\mu=1}^{Q-2} \frac{1}{2 \sin \frac{\pi \mu}{2(Q-1)}} \\
&\quad - \frac{Q^{1/2}(q^{1/2}-1)}{Q-1} \sum_{k=1}^{N-1} \frac{1}{2 \sin \frac{\pi k}{2N}}.
\end{aligned}$$

When  $r = 1$ , i.e.  $Q = q$  and  $N = 1$ , by Lemma 6 we get

$$\begin{aligned}
|\Gamma_s| &\leq 1 + \frac{q}{2(q-1)} \left[ 2(q-1) \left( \frac{\ln(q-1)}{\pi} + 0.263 \right) \right. \\
&\quad \left. + \frac{3\pi}{8(q-1)} \right] \\
&\leq 2 + \left( \frac{m \ln 2}{\pi} + 0.263 \right) 2^m.
\end{aligned}$$

When  $r > 1$ , by Lemma 6 we have

$$|\Gamma_s| \leq 1 + \frac{Q^{1/2}q^{1/2}}{2(Q-1)} \left[ 2(Q-1) \left( \frac{\ln(Q-1)}{\pi} + 0.263 \right) \right]$$

$$\begin{aligned}
& + \frac{3\pi}{8(Q-1)} \Big] - \frac{Q^{1/2}(q^{1/2}-1)}{2(Q-1)} 2N \left( \frac{\ln N}{\pi} + 0.163 \right) \\
& < 2 + \left( \frac{\ln Q}{\pi} + 0.263 \right) Q^{1/2} q^{1/2} \\
& \quad - \left( \frac{\ln N}{\pi} + 0.163 \right) \frac{Q^{1/2}}{q^{1/2}+1} \\
& \approx \left[ \frac{(n-m)\ln 2}{\pi} + 0.263 \right] 2^{n/2} \\
& \quad - \left[ \frac{(n-2m)\ln 2}{\pi} + 0.163 \right] 2^{n/2-m} + 2.
\end{aligned}$$

Hence for any  $r \geq 1$  approximately we can write that

$$\begin{aligned}
|\Gamma_s| & \leq \left[ \frac{(n-m)\ln 2}{\pi} + 0.263 \right] 2^{n/2} \\
& \quad - \left[ \frac{(n-2m)\ln 2}{\pi} + 0.163 \right] 2^{n/2-m} + 2.
\end{aligned}$$

The following lemma is an equivalent formulation of [10, Theorem 5].

**Lemma 9.** [10] *Let  $h$  be the Carlet-Feng function of  $k$  variables. Then for any  $a \in \mathbb{F}_{2^k}$ ,*

$$|W_h(a)| \leq \left( \frac{k \ln 2}{\pi} + 0.485 \right) 2^{k/2+1}.$$

**Theorem 6.** *Let  $F$  be the Boolean function defined in Construction 2. Then*

$$\begin{aligned}
\mathcal{N}_F & \geq 2^{n-1} - \left[ \frac{(n-m)\ln 2}{\pi} + 0.263 \right] 2^{n/2} \\
& \quad - \left[ \frac{(n-m)\ln 2}{\pi} + 0.485 \right] 2^{(n-m)/2} \\
& \quad + \left[ \frac{(n-2m)\ln 2}{\pi} + 0.163 \right] 2^{n/2-m} - 2.
\end{aligned}$$

*Proof:* We compute  $W_F(a, b)$  for any  $(a, b) \in \mathbb{F}_Q \times \mathbb{F}_q$ . When  $(a, b) = (0, 0)$ , we have  $W_F(a, b) = 0$  since  $F$  is balanced. When  $(a, b) \neq (0, 0)$ , we have

$$\begin{aligned}
W_F(a, b) & = -2 \sum_{(x, y) \in \text{supp}(F)} (-1)^{\text{Tr}(ax) + \text{tr}(by)} \\
& = -2 \sum_{\gamma \in \Delta_s} \sum_{y \in \mathbb{F}_q^*} (-1)^{\text{Tr}(a\gamma y^u) + \text{tr}(by)} \\
& \quad - 2 \sum_{x \in \Delta_l} (-1)^{\text{Tr}(ax)}.
\end{aligned}$$

If  $a = 0$ ,  $b \neq 0$ , then

$$\begin{aligned}
W_F(a, b) & = -2 \sum_{\gamma \in \Delta_s} \sum_{y \in \mathbb{F}_q^*} (-1)^{\text{tr}(by)} - 2 \sum_{x \in \Delta_l} 1 \\
& = -2 \times \frac{Q}{2} \times (-1) - 2 \times \frac{Q}{2} \\
& = 0.
\end{aligned}$$

If  $a \neq 0$ ,  $b = 0$ , then

$$W_F(a, b)$$

$$\begin{aligned}
& = -2 \sum_{\gamma \in \Delta_s} \sum_{y \in \mathbb{F}_q^*} (-1)^{\text{Tr}(a\gamma y^u)} - 2 \sum_{x \in \Delta_l} (-1)^{\text{Tr}(ax)} \\
& = -2 \sum_{\gamma \in \Delta_{s'}} \sum_{y \in \mathbb{F}_q^*} (-1)^{\text{Tr}(\gamma y)} - 2 \sum_{x \in \Delta_l} (-1)^{\text{Tr}(ax)} \\
& \quad (\text{note that } \alpha^{s'} = a\alpha^s) \\
& = -2\Lambda_{s'} + W_\omega(a),
\end{aligned}$$

which leads to

$$\begin{aligned}
& |W_F(a, b)| \\
& \leq \begin{cases} \left[ \frac{(n-m)\ln 2}{\pi} + 0.485 \right] 2^{(n-m)/2+1} + 2^m & \text{if } r = 1 \\ \left[ \frac{(2n-3m)\ln 2}{\pi} + 0.748 \right] 2^{(n-m)/2+1} + 2^m + 2 & \text{if } r > 1. \end{cases}
\end{aligned}$$

according to Lemma 7 and Lemma 9. If  $ab \neq 0$ , it is easy to see that

$$\begin{aligned}
W_F(a, b) & = -2 \sum_{\gamma \in \Delta_s} \sum_{y \in \mathbb{F}_q^*} (-1)^{\text{Tr}(b^{-u}a\gamma y) + \text{tr}(y^{\tilde{u}})} \\
& \quad - 2 \sum_{x \in \Delta_l} (-1)^{\text{Tr}(ax)} \\
& = -2\Gamma_{s'} + W_\omega(a)
\end{aligned}$$

for some  $0 \leq s' \leq Q-2$ , where  $\tilde{u}u \equiv 1 \pmod{q-1}$ . Then Lemma 8 and Lemma 9 implies that

$$\begin{aligned}
|W_F(a, b)| & \leq 2 \left[ \frac{(n-m)\ln 2}{\pi} + 0.263 \right] 2^{n/2} \\
& \quad + 2 \left[ \frac{(n-m)\ln 2}{\pi} + 0.485 \right] 2^{(n-m)/2} \\
& \quad - 2 \left[ \frac{(n-2m)\ln 2}{\pi} + 0.163 \right] 2^{n/2-m} + 4.
\end{aligned}$$

Therefore, we finally get that

$$\begin{aligned}
& \max_{(a, b) \in \mathbb{F}_Q \times \mathbb{F}_q} |W_F(a, b)| \\
& = \max \left\{ \max_{a \in \mathbb{F}_Q^*} |W_F(a, 0)|, \max_{(a, b) \in \mathbb{F}_Q^* \times \mathbb{F}_q^*} |W_F(a, b)| \right\} \\
& \leq \left[ \frac{(n-m)\ln 2}{\pi} + 0.263 \right] 2^{n/2+1} \\
& \quad + \left[ \frac{(n-m)\ln 2}{\pi} + 0.485 \right] 2^{(n-m)/2+1} \\
& \quad - \left[ \frac{(n-2m)\ln 2}{\pi} + 0.163 \right] 2^{n/2-m+1} + 4.
\end{aligned}$$

Then we can complete the proof applying the relation

$$\mathcal{N}_F = 2^{n-1} - \frac{1}{2} \max_{(a, b) \in \mathbb{F}_Q \times \mathbb{F}_q} |W_F(a, b)|.$$

It can be seen from the expression of the lower bound of the nonlinearity of  $F$  given in Theorem 6 that, for a fixed  $n$ , the bigger  $m$  is, the higher the lower bound is. In particular, when  $m = n/2$ , this lower bound is higher than the one proposed in

[8] and even higher than the one proposed in [14] when  $n \geq 12$ . See Table I for the comparison of lower bounds obtained in Theorem 6 and some known ones for some values of  $n$  in this case.

For small values of number of variables, we compute the exact values of the nonlinearity of  $F$  for certain choices of  $u$  (from different cyclotomic cosets modulo  $(2^m - 1)$ ). Since the computational results for the case  $r = 1$  have already been presented in [8], we need only to focus on the case  $r > 1$  here. Several results for the case  $r = 3$  are listed in Table II. By comparing these values with nonlinearity of the Carlet-Feng functions and the functions constructed in [14] in the corresponding cases, it can be seen that, at least for these numbers of variables, nonlinearity of functions from Construction 2 is high.

#### D. Immunity against FAA's

As indicated in [1], when  $r = 1$  and  $u = 2^t$ , the function  $F$  in Construction 2, which can be viewed as a variant of a balanced Tu-Deng function, behaves almost worst against FAA's. The reason is that  $F(x, y)$  only differs from  $f(x, y)$ , the function defined in Construction 1, when  $y = 0$ , so for any linear function  $L(y)$  of  $m$  variables, we have  $L(y)F(x, y) = L(y)f(x, y)$ , which implies  $\deg LF \leq m + 1$  since in this case  $\deg f = m = n/2$ . When  $r > 1$ , a similar argument shows that, for any linear function  $L(y)$  of  $m$  variables,  $\deg LF \leq \deg f + 1$ . According to Theorem 1, the degree of  $f$  is  $n - m$ . Hence we are clear that, for a fixed  $n$ , the smaller  $r$  is (or the bigger  $m$  is), the worse behavior the functions from Construction 2 against FAA's have, when  $u = 2^t$ . For the case  $u \neq 2^t$ , the behavior of functions from Construction 2 against FAA's varies, and it is an interesting problem to study for what choice of  $u$   $F$  will play particularly well.

Fixing  $s = l = 0$  and choosing certain values of the parameters  $r, m, u$  (from different cyclotomic cosets modulo  $(2^m - 1)$ ), we do some computer experiments to observe whether the pair  $(e, d)$  with  $e < n/2$  and  $e + d < n$  such that there is a function  $h$  satisfying  $\deg h \leq e$  and  $\deg hF \leq d$  exists. It turns out that:

- (1) in the cases  $r = 3, m = 3$ , (i.e.  $n = 12$ ), such pair with  $e + d \leq n - 2$  does not exist for any possible  $u$ ;
- (2) in the case  $r = 3, m = 4$ , (i.e.  $n = 16$ ), such pair with  $e + d \leq n - 2$  does not exist for any possible  $u$ ;
- (3) in the case  $r = 5, m = 3$ , (i.e.  $n = 18$ ), such pair with  $e + d \leq n - 2$  does not exist, and the pairs  $(3, 14)$  and  $(4, 13)$  ( $e + d = n - 1$ ) do not exist, for any possible  $u$ ;
- (4) in the case  $r = 3, m = 5$ , (i.e.  $n = 20$ ), such pair with  $e + d \leq n - 2$  does not exist for any possible  $u$  except 1, and the pairs  $(1, 15)$ ,  $(2, 15)$ ,  $(3, 15)$  and  $(4, 14)$  do not exist for  $u = 1$ . Besides, the pair  $(4, 15)$  ( $e + d = n - 1$ ) does not exist for  $u = 11$ . These experimental results imply that the function  $F$  has good immunity against FAA's.

## VI. CONCLUSION AND FURTHER WORK

We propose a general approach to construct Boolean functions with good cryptographic properties based on decompo-

sitions of additive groups of finite fields. A class of balanced functions with high nonlinearity and optimal algebraic degree are constructed via this approach. Algebraic immunity of these functions is optimal provided a more generalized combinatorial conjecture on binary strings is true, and immunity of them against fast algebraic attacks is also good according to some computational results. This class of functions covers some known classes of functions with (potential) optimal algebraic immunity constructed based on additive decompositions of finite fields.

Finally we should point out that, when  $r = 1$ , behavior of the function  $F$  in Construction 2 against FAA's was theoretically studied in [11]. Therefore, when  $r > 1$ , how to study behavior of  $F$  against FAA's theoretically will be a further research topic of the authors.

## REFERENCES

- [1] C. Carlet, "On a weakness of the Tu-Deng function and its repair," *Cryptology ePrint Archive*, report 2009/606, 2009.
- [2] C. Carlet, "Boolean functions for cryptography and error correcting codes," In *Monography Boolean Methods and Models*, London, England: Cambridge University Press, 2010.
- [3] C. Carlet, K. Feng, "An infinite class of balanced functions with optimal algebraic immunity, good immunity to fast algebraic attacks and good nonlinearity," in *Proc. Adv. Cryptol.-ASIACRYPT08*, LNCS, Berlin, Germany: Springer-Verlag, vol. 5350, pp. 425–440, 2008.
- [4] G. Cohen, J.-P. Flori, "On a generalized combinatorial conjecture involving addition mod  $2^k - 1$ ," *Cryptology ePrint Archive*, report 2011/400, 2011.
- [5] N. Courtois, W. Meier, "Algebraic attack on stream ciphers with linear feedback," In *Proc. Adv. Cryptol.-EUROCRYPT03*, LNCS, Berlin, Germany: Springer-Verlag, vol. 2656, pp. 345–359, 2003.
- [6] N. Courtois, "Fast algebraic attacks on stream ciphers with linear feedback," In *Proc. Adv. Cryptol.-CRYPTO03*, LNCS, Berlin, Germany: Springer-Verlag, vol. 2729, pp. 176–194, 2003.
- [7] R. Graham, D. Knuth, O. Patashnik, "Concrete Mathematics," 2nd edition, Boston, US: Addison-Wesley, 1994.
- [8] Q. Jin, Z. Liu, B. Wu, X. Zhang, "A general conjecture similar to T-D conjecture and its applications in constructing Boolean functions with optimal algebraic immunity," *Cryptology ePrint Archive*, report 2011/515, 2011.
- [9] R. Lidl, H. Niederreiter, "Finite Fields," London, England: Cambridge University Press, 1997.
- [10] M. Liu, Y. Zhang, D. Lin, "Perfect algebraic immune functions," In *Proc. Adv. Cryptol.-ASIACRYPT 12*, LNCS, Berlin, Germany: Springer-Verlag, vol. 7658, pp. 172–189, 2012.
- [11] M. Liu, D. Lin, "Almost Perfect Algebraic Immune Functions with Good Nonlinearity," *Cryptology ePrint Archive*, report 2012/498, 2012.
- [12] F. MacWilliams, N. Sloane, "The Theory of Error-Correcting Codes," Amsterdam, Netherlands: North-Holland, 1977.
- [13] W. Meier, E. Pasalic, C. Carlet, "Algebraic attacks and decomposition of boolean functions," In *Proc. Adv. Cryptol.-EUROCRYPT04*, LNCS, Berlin, Germany: Springer-Verlag, vol. 3027, pp. 474–491, 2004.
- [14] D. Tang, C. Carlet, X. Tang, "Highly nonlinear Boolean functions with optimum algebraic immunity and good behavior against fast algebraic attacks," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 653–664, 2013.
- [15] Z. Tu, Y. Deng, "A conjecture about binary strings and its applications on constructing Boolean functions with optimal algebraic immunity," *Des. Codes Cryptogr.*, vol. 60, no. 1, pp. 1–14, 2011.

TABLE I  
COMPARISON OF LOWER BOUNDS OF NONLINEARITY IN THE CASE  $n = 2m$

$n$	6	8	10	12	14	16	18	20	22	24	26
LB in Th. 6	20	102	457	1930	7936	32211	129863	521671	2091509	8376484	33528475
LB in [8]	18	93	429	1858	7762	31808	128949	519628	2086991	8366580	33506919
LB in [14]	20	102	458	1929	7931	32195	129823	521577	2091288	8376003	33527429

TABLE II  
NONLINEARITY OF  $F$  IN THE CASE  $r = 3, s = l = 0$

$n$	$\mathcal{N}_F$		$\mathcal{N}_{C-F}$ in [3]	$\mathcal{N}_{T-C-T}$ in [14]	$2^{n-1} - 2^{n/2-1}$
12	$u = 1$	1982	1970	1982	1984
	$u = 6$	1964			
16	$u = 1$	32408	32530	32508	32512
	$u = 14$	32406			